

## PISA 2021 Adjusted Design

This document presents an adjusted design for PISA 2021, which takes into account that many countries will not be able to conduct any PISA field operations this year.

The proposal has been prepared by the OECD Secretariat under the guidance of the PISA Executive Group. The proposal will allow PISA to provide a baseline assessment in the aftermath of the Covid-19 crisis, which the Executive Group considers important. To this end, the assessment will also include a module with questions to students on their learning experiences during this crisis.

Although most materials in any PISA cycle are reused from previous cycles, the omission of the Field Trial (FT) will require some changes to the design for PISA 2021 to ensure that the resulting data are compliant with the PISA technical standards. The main purpose of the FT was to test new items before inclusion in the Main Survey (MS) and prepare for an adaptive version of the test of mathematics. With no FT, the proposal is to conduct PISA 2021 mostly with existing material and postpone the adaptive version of mathematics to the PISA 2024 assessment.

New items will be included in the assessment to prepare for the transition of mathematics from linear to adaptive in 2024 and obtain data on other new elements where needed. Where new items are included, they will only be used for analysis and reporting if they meet the same data quality standards that would normally have been applied during FT analyses.

The key components of the adjusted design are:

- Countries that already participated in PISA 2018 will not need to conduct a FT in 2020.
- PISA 2021 will focus on the use of existing instruments that have been proven to meet data quality standards.
- All investments made in preparation for PISA 2021 will be maintained and used either in PISA 2021 or future cycles, including the translations and adaptations already made by countries.
- PISA 2021 will give the same opportunity for international benchmarking, trend measures and analyses of contextual information as other cycles, with the exception of the innovative domain that will only be piloted in this cycle.
- There will be no increase in the overall scope of work or budgets for national PISA centres.

### Field Trial situation before school closures

The PISA 2021 FT was originally scheduled to run from 1 March through 26 September 2020. However, given the impact of Covid-19, most countries/economies have been unable to proceed with the FT as planned. For some it remains unclear if a FT will be possible in 2020 and others know that they will not be able to conduct a FT this year.

Twelve participants were able to conduct part of the FT data collection before schools were closed. The limited data from these countries will be used to identify items that do not meet the data quality standards and recommend others items to be included for the MS. However, the amount of data is

not sufficient to make the usual full selection of items for the MS or to convert the mathematics test into an adaptive test as had been planned for PISA 2021. The data will also not be sufficient to complete the feasibility study of online delivery on Chromebooks that had been planned for this cycle.

### Revised Field Trial plan

The proposal for an adjusted design takes into account that many countries/economies will not be able to schedule a FT in 2020. In the adjusted design, the FT activities will be adapted to each country's needs and circumstances.

The revised plan distinguishes between three groups of countries/economies:

#### **A. Countries/economies that participated in PISA 2018 and did not already begin FT for 2021**

- No FT will be required for countries that already participated in PISA 2018 since these countries have already implemented the majority of materials for 2021 in previous cycles and are familiar with PISA operations. Countries will instead be required to conduct a preparation exercise, which does not involve data collection in schools. The preparation exercise is done by the PISA national centre and is used to practice and gain experience in the data management and sampling procedures that are new in PISA 2021.<sup>1</sup>
- If countries wish to gain experience with PISA data collection before the MS in 2021, for example if they have a new PISA national centre (NC), they will be offered to conduct a limited FT (or “mini” FT). The International contractors will provide guidance on minimum sample sizes and timelines to NCs interested in this option.

#### **B. Countries/economies that already began the FT for 2021**

- Countries that already started their FT before school closures are asked to code open-ended responses to mathematics and creative thinking items as well as complete occupational coding for the questionnaires and submit the FT data already collected according to the existing plan.<sup>2</sup> This includes maintaining the original schedule for data submission by 31 July. International contractors will provide NCs with guidelines for completing the FT data management and sampling activities and submitting the database for data cleaning and analysis.
- These countries can then proceed to preparations for the MS without further FT activities.

#### **C. Countries/economies new to PISA or switching from Paper-Based Assessment (PBA) to Computer-Based Assessment (CBA) in PISA 2021**

---

<sup>1</sup> NC staff will be asked to follow the prescribed data management steps and complete all sampling procedures (through ST6) and submit the database according to the procedures outlined in the Data Management Manual. The proposed deadline for completing this exercise is 30 November, with the expectation that feedback from this experience will be used to guide the MS trainings for data management and sampling.

<sup>2</sup> Denmark, Ireland, Israel, Malta, Netherlands, Philippines, Qatar, United Kingdom, UK-Scotland, Sweden and Portugal.

- Countries that are new to PISA or are administering PISA as CBA for the first time are encouraged to complete a mini FT within a small number of schools (3-10) and submit data by the Batch 3 deadline (30 November) for data cleaning and analysis. If a mini FT is not possible, a preparation exercise is necessary. It should be noted that there will be more inherent risks with respect to data quality without a mini FT for participants within this category, particularly those testing in languages that have not been administered in previous cycles.

### Main Survey plan

The MS for PISA 2021 will be maintained within the original timeframe. The first countries/economies are scheduled to test in March 2021. Countries that need to reschedule their MS testing dates may be able to do so in consultation with the PISA international contractors. The sample sizes for the MS is expected to remain unchanged.

### Transition to adaptive test

The plan to make all three tests of the core domains in PISA adaptive by 2024 will be maintained with an adjusted transition plan. The original transition plan was to transform one domain per cycle from linear to adaptive, starting with reading in 2018, adding mathematics in 2021, and then science in 2024.

The change of a test from linear to adaptive requires stable difficulty parameters for all units. The data from the MS in 2021 will be used to transform the mathematics test from linear to adaptive in 2024. This means that under this plan, mathematics will be made adaptive in 2024 at the same time as the science test.

The tables below show the original and new transition plan for adaptive testing.

<i><b>Original transition plan</b></i>			
	Reading	Maths	Science
<b>2018</b>	Adaptive		
<b>2021</b>	Adaptive	Adaptive	
<b>2024</b>	Adaptive	Adaptive	Adaptive
<b>Beyond</b>	Adaptive	Adaptive	Adaptive

<i><b>New transition plan</b></i>			
	Reading	Maths	Science
<b>2018</b>	Adaptive		
<b>2021</b>	Adaptive		
<b>2024</b>	Adaptive	Adaptive	Adaptive
<b>Beyond</b>	Adaptive	Adaptive	Adaptive

### Assessment components

The assessment components for PISA 2021 will to a large extent remain unchanged. All three core domain tests of reading, mathematics and science will be included, with mathematics as the major domain as planned. However, some important changes will be made due to the limited FT. Most importantly, a larger item pool of the newly developed mathematics items as well as newly developed contextual questions will be included in the MS with the expectation that some items will be deleted during analysis, as would have typically happened after the FT. Following analysis, only items that meet psychometric standards will be used for analyses and reporting. The same applies for creative thinking which would be included on a pilot basis. Thus, some of the analysis that would previously apply to the FT will now be embedded in the MS analysis.

The table below shows, for each component of PISA 2021, whether there will be changes with the adjusted design.

Component	Adjustment compared with original PISA 2021 design
<b>Reading</b>	This adjustment will not have an impact on the reading test, which will be administered as planned through a reduced adaptive design from 2018.
<b>Mathematics</b>	<p>This adjustment will not have an impact on the trend mathematics items. The exception is that these materials will be administered through intact clusters.</p> <p>Newly developed items for mathematics will be included in the MS and following analysis, only those items that meet psychometric standards will be included in analyses and reporting.</p>
<b>Science</b>	This adjustment will not have an impact on the science test, which will be administered as it was in 2018.
<b>Creative thinking</b>	<p>The creative thinking items will be included in the MS on a pilot basis. Selection of the items will be based on analyses of the available FT data and previous pilot studies. Decisions about the reporting of the results, further development and administration of the test will be taken based on the data collected in the pilot.</p> <p>Given that the results of the creative thinking pilot will only be known in 2022, there would be very little time to make improvements in the domain for administration in 2024. In order to maximise time and resources invested in the development of the innovative domains, it is proposed to proceed with Learning in the Digital World as the innovative domain for PISA 2024 and offer an improved version of creative thinking, using the results of the pilot study, in a successive cycle.</p>
<b>Financial literacy</b>	This adjustment will not have an impact on the trend items for financial literacy. Five new items developed for PISA 2021 will be included in the MS, and following analysis, only those items that meet psychometric standards will be included in analyses and reporting.
<b>Student and school questionnaires</b>	<p>The student and school questionnaires will be assembled from the 2021 FT questionnaire, which consist of both trend and newly developed questions.</p> <p>In addition, PISA 2021 will offer new questions focusing on the global pandemic totalling approximately 5 minutes as part of the 35-minute Student Questionnaire. A small number of new questions around this topic will be considered for inclusion in the School Questionnaire. We elaborate in more detail on this new module toward the end of this document. In all cases, reporting on new questions will be conditional on MS data passing crucial quality checks. Any scaled indices would have to be confirmed based on MS data as well.</p>
<b>Optional questionnaires</b>	The following optional questionnaires will be offered as planned: a student <i>Well-Being questionnaire</i> , a student <i>Information and Communication Technology Familiarity questionnaire</i> , a student <i>Financial Literacy questionnaire</i> , a <i>Parent questionnaire</i> , and a <i>Teacher questionnaire</i> organized into two sections: a 25-minute regular teacher

	questionnaire and a 15-minute teacher well-being questionnaire. In all cases, reporting on new questions included as part of these optional questionnaires will be conditional on MS data passing crucial quality checks. Any scaled indices would have to be confirmed based on MS data as well.
--	---

### New questionnaire module on global crises

A new questionnaire module on the current global crisis will be developed. How did students spend their time during school closure? What kind of support did they receive from their schools and parents during school closures?

The global crisis is likely to impact student performance and well-being. The new module provides opportunities to better identify and understand underlying influences, which in turn speak to the extent to which additional support and resources are required within individual education systems. Analyses based on this module can contribute to strengthen education systems' preparedness for other future crises.

A brief module on the global crisis can be offered as a core part of the student questionnaire. A small number of new questions around this topic will be considered for inclusion in the school questionnaire.

The PGB raised an important point regarding the need for sensitivity in the development of the new questionnaire module concerning how students and schools coped with the crisis.

In order to be sensitive to the potential impact of the current global crisis (illness, loss of family member, economic distress, etc.), the following three approaches can be adopted.

- First, individual questionnaire items in the new module will be linked to the PISA 2021 questionnaire framework. This clarifies their measurement objectives and their analytical application. This will guide the development and revision of the questionnaire items to ensure that they measure targeted constructs while maintaining awareness of respondent sensitivities.
- Second, the International Consortium is ready to contribute their time and expertise to conduct a small-scale virtual cognitive lab. The results should identify how respondents react to and perceive questions focusing on the impact of the global health crisis on their learning opportunities and experiences. The items, then, can be refined, elaborated, or rephrased as necessary. Countries are invited to volunteer their students to participate in the cognitive lab.
- Third, countries are invited to review and comment on all new questionnaire items for sensitivity. Feedback from countries will be considered for the finalisation and selection of items for the MS.

The second and third approaches are possible only if the development of the new module starts by mid-April. The final master version of the 2021 MS questionnaire items has to be ready not later than mid-July, in order to have the final national version of the MS questionnaire ready by November.

The inclusion of the new module in PISA 2021 MS will bear translation and verification costs. The Secretariat will share further information regarding cost estimates and how these will be covered.

## Data quality standards in PISA 2021 adjusted design

### *Cognitive tests*

Under this adjusted design, PISA 2021 will be subject to the same data quality standards as other PISA cycles. In particular, after the MS data collection, analyses at the test and domain level will be conducted to verify the data yield and quality. Additional checks will be conducted on any new material to replicate FT analysis and select the portion of new material that is fit for reporting.

Reading and science: Data for the trend domains of reading and science will be analysed in combination with PISA 2018 data to detect inconsistencies and deviations. Where inconsistencies in response patterns are detected, the reasons will be investigated in relation to the content of the items or their format. As in past PISA cycles, isolated deviations from expected patterns would be treated by assigning cycle-and-country specific item parameters during scaling, or by dropping the items from scaling. The additional analyses related to the scaling of reading fluency items outlined in the original FT analysis plan will be conducted with the MS data, and will inform the scaling approach for PISA 2021.

Mathematics: For mathematics, this design is expected to rely on all trend mathematics clusters from 2018 and as many as possible of the new mathematics items from the FT. This content would be organised according to a balanced incomplete block design to result in hour-long mathematics tests with students taking a combination of trend and new materials. The 2021 integrated design is expected in many ways to follow the 2018 integrated design with mathematics as the major domain. The total sample size is expected to be maintained to (or very close to) the current 6,300 assessed students. For new mathematics items, the first set of analyses will involve descriptive response-frequency analyses, timing analyses, coding-reliability analyses and preliminary scale-fit analysis at the national and international level. These analyses will be conducted in batches (first on early-testing countries/economies) and will help select the mathematics items to be retained for reporting. The usual checks applied after the FT will be applied at this stage; open-ended items that show poor inter-rater reliability, items that show high levels of non-response or long response times, items that show poor fit to the mathematics scale (e.g. low item-rest correlations, low discrimination) or local dependencies, and items with problematic IRT characteristics will be considered for deletion. It is expected that some new items will be discarded at this stage, similar to FT analysis. Once the item selection is completed an international scale will be constructed based on the remaining item pool. Dimensionality analyses will be conducted to confirm the ability to report trend and new mathematics units on a single scale (if multi-dimensionality is detected, two mathematics scales will be developed – one representing the trend construct, similar to 2018, and one representing the new framework, mostly based on new items). Item-by-country/language interactions will be further investigated, and violations of invariance across country/language groups or (for trend units) over time will result in unique parameters (within a partial-invariance model) or in isolated item omissions from scaling, similar to reading in 2018. Based on the scaling results, a described proficiency scale will be developed for the new mathematics framework, and the ability to report subscale performance in mathematics will be investigated.

Creative thinking: Evaluation of quality of the items including analysis of inter-rater reliability and invariance of the item parameters will be performed in a first phase. Results of this analysis will be presented to countries with a proposed plan on the reporting of the pilot results, further development and administration of the test.

## **Data quality standards in PISA 2021 adjusted design**

### *Questionnaires*

The questionnaire analysis plan for MS data under the new design will incorporate key elements of the 2021 FT analysis plan for questionnaires in order to perform additional quality checks for new items that are typically performed during the FT. The analyses will involve five phases: (1.) descriptive response frequency analyses, including timing analyses, and analyses of student response styles; (2.) Dimensionality analyses; (3.) IRT scaling; (4.) Scaling of ESCS; and (5.) Associations between questionnaire constructs and with achievement proxies. These analyses will be conducted in batches (first on early-testing countries/economies) and will help select the questionnaire items and derived variables to be retained for reporting. The checks originally proposed for after the FT will be applied at this stage. The following criteria will be used to flag items or scaled indices for potential exclusion from reporting: high missing response rates across multiple countries; unusually long response times across multiple countries; poor loadings of items on the intended constructs; poor IRT parameters (discrimination, difficulty), misfit, or lack of measurement invariance for items developed as part of scaled indices; and low reliability and high degrees of IRT misfit across multiple countries for scaled indices.